wwPDB X-ray Structure Validation Summary Report (i)

Nov 15, 2023 – 04:08 PM JST

PDB ID	:	6JUY
Title	:	Crystal Structure of ArgZ, apo structure, an Arginine Dihydrolase from the Ornithine-Ammonia Cycle in Cyanobacteria
Authors	:	Zhuang, N.; Li, L.; Wu, X.; Zhang, Y.
Deposited on	:	2019-04-15
Resolution	:	2.97 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

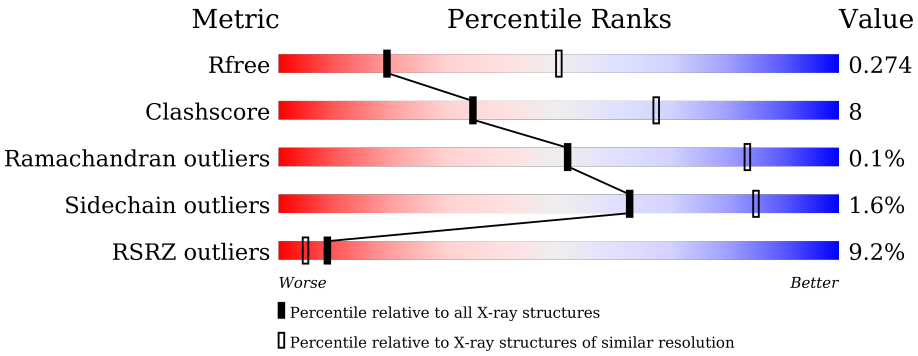
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.97 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2754 (3.00-2.96)
Clashscore	141614	3103 (3.00-2.96)
Ramachandran outliers	138981	2993 (3.00-2.96)
Sidechain outliers	138945	2996 (3.00-2.96)
RSRZ outliers	127900	2644 (3.00-2.96)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	707	6% 74% 16% 10%
1	B	707	7% 73% 17% • 10%
1	C	707	7% 74% 19% 7%
1	D	707	13% 70% 17% • 12%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 18519 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Sll1336 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	639	Total	C	N	O	S	0	0	0
			4736	3005	812	889	30			
1	B	637	Total	C	N	O	S	0	0	0
			4722	2991	817	885	29			
1	C	656	Total	C	N	O	S	0	0	0
			4723	3000	803	893	27			
1	D	623	Total	C	N	O	S	0	0	0
			4333	2745	737	823	28			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP P74535
A	0	ALA	-	expression tag	UNP P74535
B	-1	GLY	-	expression tag	UNP P74535
B	0	ALA	-	expression tag	UNP P74535
C	-1	GLY	-	expression tag	UNP P74535
C	0	ALA	-	expression tag	UNP P74535
D	-1	GLY	-	expression tag	UNP P74535
D	0	ALA	-	expression tag	UNP P74535

- Molecule 2 is water.

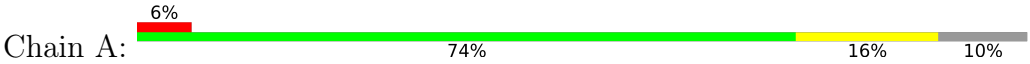
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	5	Total	O	0	0
			5	5		

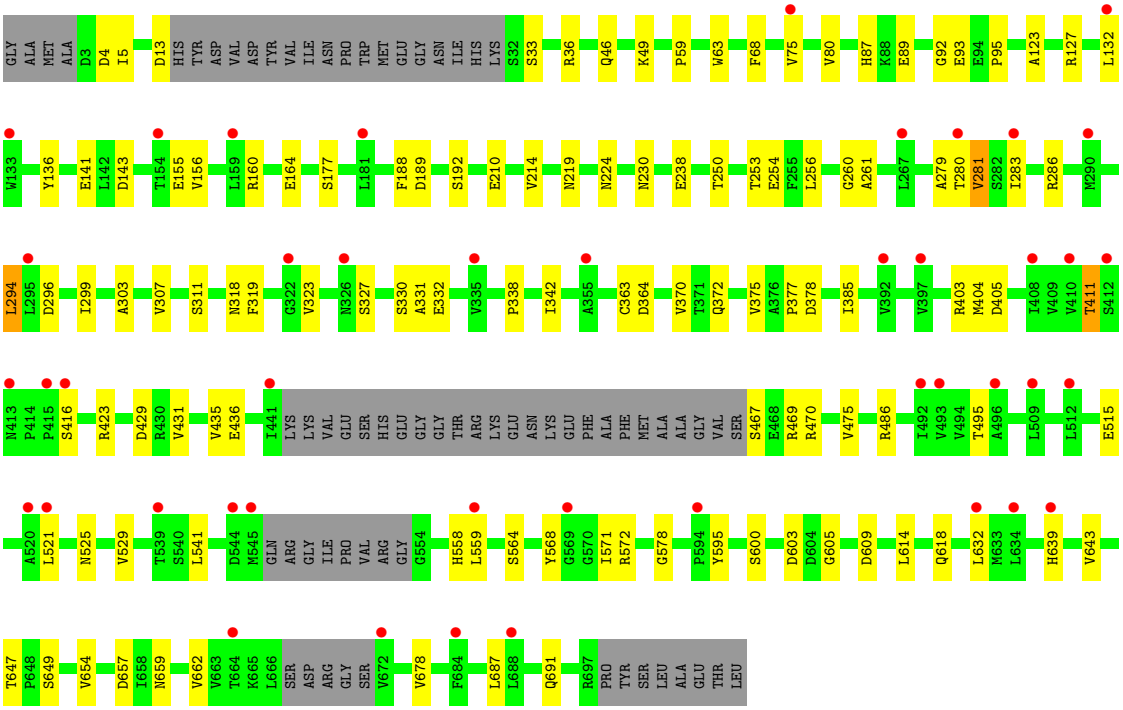
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: Sll1336 protein

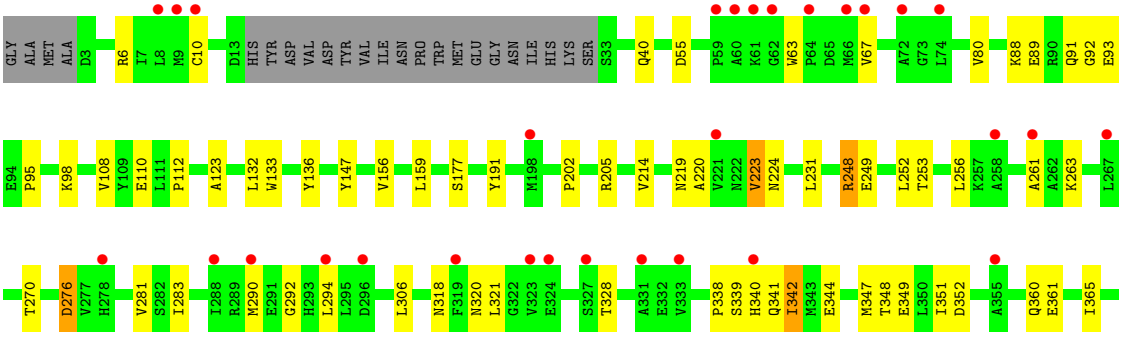
Chain A:

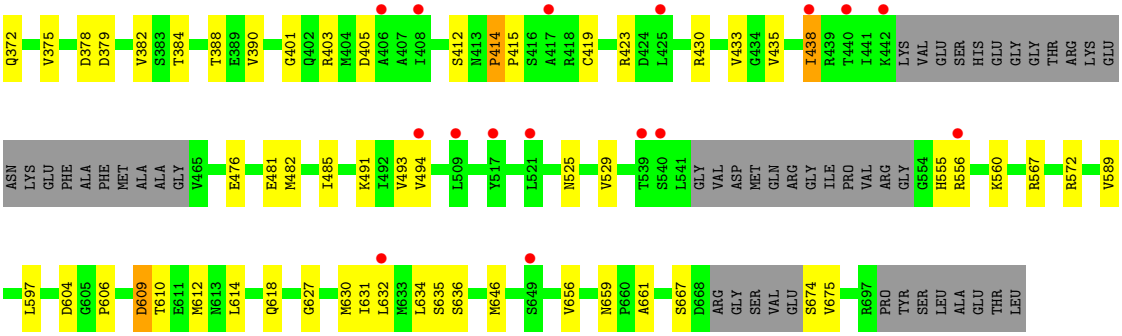




• Molecule 1: Sll1336 protein

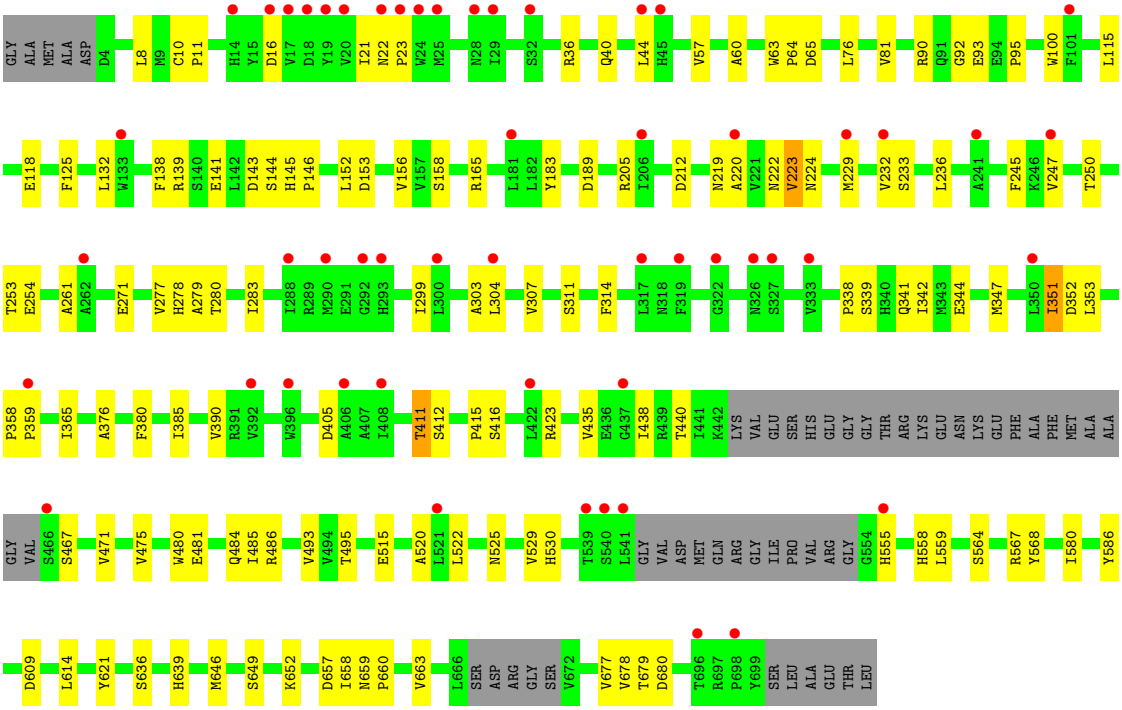
Chain B:





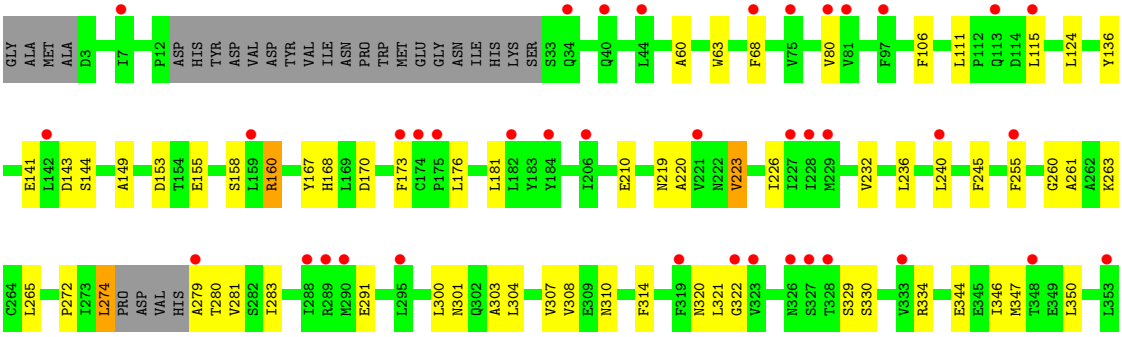
• Molecule 1: Sll1336 protein

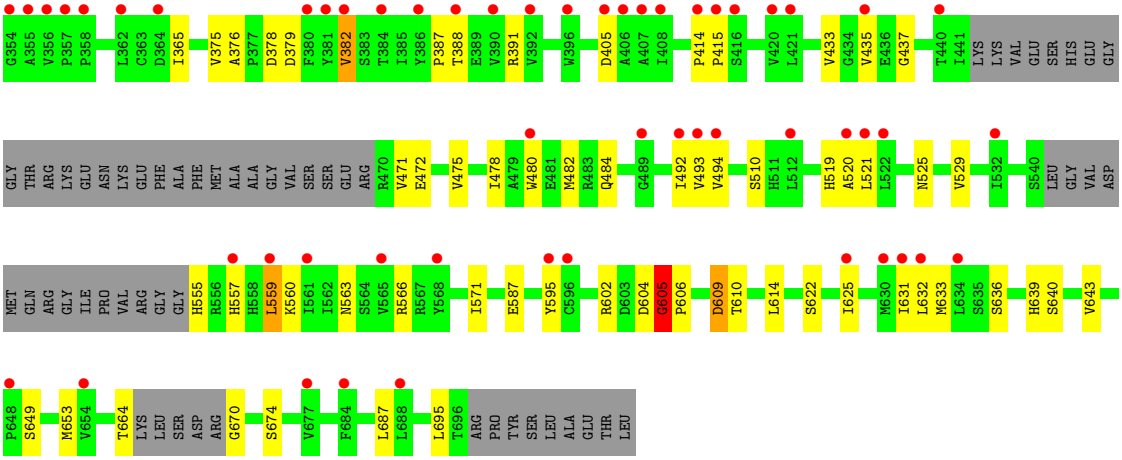
Chain C:



• Molecule 1: Sll1336 protein

Chain D:




G354 A355 V356 P357 P358 L362 C363 D364 I365 V375 A376 P377 D378 D379 F380 Y381 V382 S383 T384 I385 Y386 P387 T388 E389 V390 R391 V392 W396 D405 A406 A407 I408 P414 P415 S416 V420 L421 V433 G434 V435 E436 G437 T440 I441 LYS LYS VAL GLU SER HIS GLU GLY
GLY THR ARG LYS GLU ASN LYS GLU PHE ALA PHE MET ALA ALA GLY VAL SER SER GLU ARG R470 V471 E472 V475 I478 A479 W480 E481 M482 R483 Q484 G489 I492 V493 V494 S510 H511 L512 H519 A520 L521 L522 N525 V529 I532 S540 LEU GLY VAL ASP
MET GLN ARG GLY ILE PRO VAL ARG GLY GLY H555 R556 H557 H558 L559 K560 I561 I562 N563 S564 V565 R566 R567 Y568 I571 E587 Y595 C596 R602 D603 D604 G605 P606 D609 T610 L614 S622 I625 M630 I631 L632 M633 L634 S635 S636 H639 S640 V643
P648 S649 M653 V654 T664 LYS LEU SER ASP ARG G670 S674 V677 F684 L687 L688 L695 T696 ARG PRO TYR SER LEU ALA GLU THR LEU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	189.87Å 105.44Å 165.91Å 90.00° 113.18° 90.00°	Depositor
Resolution (Å)	41.06 – 2.97 48.74 – 2.97	Depositor EDS
% Data completeness (in resolution range)	97.1 (41.06-2.97) 97.2 (48.74-2.97)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.96 (at 2.96Å)	Xtriage
Refinement program	PHENIX 1.17_3644	Depositor
R, R_{free}	0.234 , 0.275 0.234 , 0.274	Depositor DCC
R_{free} test set	1990 reflections (3.29%)	wwPDB-VP
Wilson B-factor (Å^2)	109.2	Xtriage
Anisotropy	0.456	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 79.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.000 for 1/2*h-3/2*k,-1/2*h-1/2*k,-1/2*h+1/2*k-l 0.003 for 1/2*h+3/2*k,1/2*h-1/2*k,-1/2*h-1/2*k-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	18519	wwPDB-VP
Average B, all atoms (Å^2)	109.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.46% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	A	0.27	0/4825	0.50	0/6572
1	B	0.28	0/4814	0.51	1/6565 (0.0%)
1	C	0.32	1/4821 (0.0%)	0.54	2/6597 (0.0%)
1	D	0.27	0/4412	0.53	1/6041 (0.0%)
All	All	0.29	1/18872 (0.0%)	0.52	4/25775 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	2
1	C	0	2
1	D	0	1
All	All	0	6

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	341	GLN	CD-OE1	-5.59	1.11	1.24

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	559	LEU	CB-CG-CD1	-6.85	99.36	111.00
1	C	659	ASN	C-N-CD	-6.81	105.61	120.60
1	B	414	PRO	C-N-CD	-5.40	108.73	120.60
1	C	659	ASN	C-N-CA	5.24	144.00	122.00

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	605	GLY	Peptide
1	B	223	VAL	Peptide
1	B	634	LEU	Peptide
1	C	223	VAL	Peptide
1	C	277	VAL	Peptide

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4736	0	4513	67	2
1	B	4722	0	4476	80	1
1	C	4723	0	4314	88	1
1	D	4333	0	3874	82	0
2	A	5	0	0	0	0
All	All	18519	0	17177	297	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 8.

The worst 5 of 297 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:659:ASN:HB3	1:A:662:VAL:HB	1.55	0.88
1:D:493:VAL:HG12	1:D:520:ALA:HB3	1.55	0.88
1:C:76:LEU:HB3	1:C:152:LEU:HD12	1.59	0.82
1:B:338:PRO:HD2	1:B:342:ILE:HD11	1.63	0.80
1:B:340:HIS:O	1:B:341:GLN:NE2	2.13	0.79

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:279:ALA:O	1:C:412:SER:OG[1_565]	2.08	0.12

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:238:GLU:OE2	1:B:248:ARG:NH2[3_455]	2.13	0.07

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	629/707 (89%)	594 (94%)	34 (5%)	1 (0%)	47 80
1	B	627/707 (89%)	600 (96%)	27 (4%)	0	100 100
1	C	648/707 (92%)	619 (96%)	28 (4%)	1 (0%)	47 80
1	D	611/707 (86%)	580 (95%)	30 (5%)	1 (0%)	47 80
All	All	2515/2828 (89%)	2393 (95%)	119 (5%)	3 (0%)	51 83

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	278	HIS
1	D	605	GLY
1	A	578	GLY

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	483/601 (80%)	475 (98%)	8 (2%)	60 84
1	B	478/601 (80%)	470 (98%)	8 (2%)	60 84

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	C	452/601 (75%)	446 (99%)	6 (1%)	69 88
1	D	398/601 (66%)	391 (98%)	7 (2%)	59 83
All	All	1811/2404 (75%)	1782 (98%)	29 (2%)	62 85

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	636	SER
1	D	609	ASP
1	C	280	THR
1	D	274	LEU
1	C	271	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	278	HIS
1	D	502	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	639/707 (90%)	0.41	44 (6%)	16	9	68, 97, 133, 156	0
1	B	637/707 (90%)	0.56	46 (7%)	15	8	68, 96, 132, 185	0
1	C	656/707 (92%)	0.49	53 (8%)	12	6	74, 116, 146, 185	0
1	D	623/707 (88%)	0.70	93 (14%)	2	1	85, 125, 151, 176	0
All	All	2555/2828 (90%)	0.54	236 (9%)	9	5	68, 108, 144, 185	0

The worst 5 of 236 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	60	ALA	7.0
1	B	323	VAL	6.9
1	D	288	ILE	6.1
1	B	324	GLU	5.9
1	C	19	TYR	5.7

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.